

Choukroune
 Nouredine
 SRI

Le 24/06/2012

Classification et Analyse des données

durée : 1H : 30

Question de cours : (3 points)

1. Pour quel type de méthode de classification peut-on appliquer la mesure de cohésivité?
2. Nous avons utilisé l'algorithme BIRCH pour classifier un ensemble de données avec un rayon Ft. Les résultats obtenus montrent que les classes sont denses. Comment peut-on modifier l'algorithme pour avoir une meilleure classification avec le même rayon Ft ?

Exercice 1 : (3 points)

Nous avons réalisé un ACP sur des variables centrés réduites ; les variables des 7 premières composantes principales sont :

1	2	3	4	5	6	7
7.3154	1.9487	-0.8457	1.2547	2.6548	0.5428	0.3547

Combien de composantes principales faut-il garder à priori ? Quelle est le pourcentage de l'inertie totale expliquée par le sous-espace principal correspondant ?

Exercice 2 : (7 points)

On observe le tableau suivant :

X	64	62	66	65	61	68	70	69	67	63
Y	68	65	68	69	67	71	72	72	70	67

1. Calculer la moyenne, l'écart type et la variance de la variable X et de la variable Y.
2. Calculer l'équation de la droite de régression linéaire en x de cette série.
3. Calculer le R^2 de cette série. Commenter.
4. En inversant les rôles de x et y, calculer l'équation de la droite de régression linéaire en y de cette série.
5. Quel est leur point d'intersection ?

Exercice 3 : (7 points)

Nous avons appliqué une classification supervisée sur un ensemble de données en 4 classes, le tableau suivant représente les résultats obtenus :

Chaque ligne correspond au nombre de données qui appartient à la classe.
 Chaque colonne correspond au nombre de données classés dans la classe.

	C1	C2	C3	C4
C1	46	4	3	2
C2	5	68	5	3
C3	1	3	49	5
C4	2	4	11	52

1. Calculer la mesure Ft de cette classification.